Referee 2

Many thanks for these constructive suggestions. Our responses are in red below.

General comments

The manuscript describes a new coupled atmosphere-ocean model in a rather compact manner focusing on the parameter tuning. The model should be very useful for the community and would promise contribution to the scientific advancement in this field. This aspect is enhanced particularly by the release of the model. The description of such model with potentially broad application is useful and deserves to be published in the GMD. The current description is, however, unsatisfactory in its current form for the reasons listed below.

Major comments

1. Quantitative (and physical, in some cases) meaning of some variables discussed in the manuscript is unclear for the non-GENIE users unless the reader consults with the multiple previous papers. I do not expect that all variables are explained in details, but the highlighted variables such as “acllwr”, “albseamax”, “qthresh” “ADRAG”, and “SCF” need to be expressed with equations. Otherwise, the discussion on the presented values does not mean much for many potential readers. The term “sea ice diffusivity” is also unclear although I would imagine this represents the diffusive effect of unresolved ocean currents (and other dynamical effect) on the sea ice concentration.

We have greatly expanded the descriptions of parameters. Section 4.1

“Parameters from modules other than GOLDSTEIN were all fixed. However, some were changed from their default values on the basis of exploratory simulations:

i) The uncertain effect of clouds on long wave radiation is controlled through the dependence of cloud emissivity A on the mass absorption factor k “acllwr”, following Slingo and Slingo (1991):

$$A = 1 - e^{-\beta k W}$$

where $\beta = 1.66$ is the diffusivity factor and W the cloud liquid water path. The mass absorption factor was found to exert the strongest control on surface air temperature of the 22 key model parameters considered in PLASIM-ENTS ensembles (Holden et al 2014). The value was increased from default $k = 0.1$ to $0.2 m^2 g^{-1}$, estimated to yield a simulated global average surface air temperature of approximately 14°C in conjunction with parameter choices (ii) to (v) below.

ii) The PLASIM parameter *albseamax* defines the latitudinal variation of ocean albedo (Holden et al 2014),

$$\alpha_s = \alpha_{s0} + 0.5\alpha_{s1}[1 - cos(2\varphi)]$$

where the ocean albedo α_s is expressed in terms of latitude φ, the albedo at the equator $\alpha_{s0} = 0.069$ and the parameter that controls latitudinal variability α_{s1}. The calculated albedo is applied to both direct and scattered radiation. A high value ($\alpha_{s1} = 0.4$) was favoured for *albseamax*, leading to cooler high latitude ocean and favouring increased Southern Ocean sea-ice and deep-water formation, which both tended to be too low with default parameters.

iii) Sea ice is transported through advection and Laplacian diffusion (Edwards and Marsh, 2005), the latter taking the place of a detailed representation of unresolved advection and rheological processes. Sea-ice diffusivity (*SID*) influences AABW formation by controlling the rate at which new ice is created, and hence the strength of brine rejection (Holden et al 2013b). A high value was favoured, again to strengthen deep-water formation, but values greater that 15,000 m^2s^{-1} were found to lead to numerical instabilities in this model and *SID* was fixed at this value.

iv) The standard PLASIM expression for the dependence of sea ice albedo α_i on surface air temperature is used

$$\alpha_i = 0.5 - 0.025T_{air}$$

where T_{air} is the surface air temperature (°C). PLASIM restricts the maximum albedo to 0.7 ($T_{air} \leq -8°C$). In PLASIM-GENIE we also restrict the minimum albedo, 0.5 ($T_{air} \geq 0°C$).

v) The PLASIM-ENTS dependency of photosynthesis on soil moisture is

$$f_2(W_s) = \{(W_s/W_s^*) - q_{th}\}/\{0.75 - q_{th}\}$$

The parameter q_{th} (*qthresh)* was set to 0.1, allowing the development of vegetation in semi-arid regions (Holden et al 2014).

The ensemble was generated using a 50x6 maximin latin hypercube design, varying six GOLDSTEIN parameters, listed in in Table 1 and varied over ranges considered to reflect the plausible range for each parameter (Holden et al 2013b and references therein). The six varied parameters are isopycnal and diapycnal diffusivities, a parameter *OP1* that controls the depth profile of diapycnal diffusivity (see below), the frictional drag coefficient (GOLDSTEIN is based upon the thermocline equations with the addition of a linear drag term in the horizontal momentum equations, Edwards et al 1998), wind stress scaling (a linear scaling of the surface wind-stress is applied to compensate for the energy dissipated by frictional drag), and an Atlantic-Pacific moisture flux adjustment.

Diapycnal diffusivity is stratification-dependent (Oliver and Edwards, 2008), given by

$$k_v = k_{v0} p_0(z)^{\gamma} \left(\frac{\Delta\rho_0(z)}{\Delta\rho(z)}\right)$$

where k_{v0} (reference diapycnal diffusivity) and γ (*OP1*) are varied across the ensemble (Table 1), $p_0(z)$is a reference profile (exponentially growing with depth and equal to 1 at 2500m), $\Delta\rho_0(z)$ a reference vertical density gradient profile and $\Delta\rho(z)$ the local simulated vertical density gradient."

2. If the APM is a flux-correction parameter, should this depend on the flow? If the model simulates the moisture flux correctly, the addition of this flux correction would lead to a wrong total moisture flux. This is no longer a correction. By design, the flow responds to this parameter "forcing". I am not sure about the physical meaning of this parameter which appears to be an important tuning exercise here.

Apologies, the parameter should have been described as a flux adjustment. This change has been made throughout. Because the simulated flux is generally different from the observationally derived values and because of uncertainty in both the true value of the flux and also its effect on the system, the parameter plays an important role as a control parameter for the model. It is not clear how it should be related dynamically to the flow.

3. I wonder why the performance of only thermohaline circulation is discussed. In the introduction, it is mentioned that "dynamic ocean feedbacks are restricted to the thermohaline circulation" in the previous EMBM coupled version of the model. Then, the simulated wind-driven circulation and its interaction with the thermohaline circula- tion must be one of the selling points of the new model. Why not discussing the wind driven circulation (and its bias)? Similarly, the sea ice plays an important role for the thermohaline circulation. Why not discussing the sea ice distribution (and its bias)?

We now consider the influence of wind-driven circulation on the thermohaline circulation with reference to high-frequency variability, and including additional Figure 6e:

"Figure 6e illustrates wind-driven AMOC variability, behaviour that is absent from GENIE-1 (Sarojini et al 2011), because it is forced with annual averaged climatological winds. The maximum Atlantic overturning circulation is plotted through an arbitrary year (year 100 of a "spin-on" simulation), together with the 100-year mean and standard deviation. "

The sea-ice bias is discussed in some detail. A plot is not included as the bias is very large so that a plot would add little explanatory value:

"Sea-ice distributions (not illustrated) exhibit a systematic bias towards low southern sea-ice area across the ensemble, with an annual average of 2.8 million km^2 in the subjective tuning; this compares to observational estimates of 11.5 million km^2 (Cavalieri et al 2003). Surface air temperature over the southern ocean is warm biased with respect to the reanalysis data, despite a modest cold bias in the global temperature (Figure 2). While this may in part be a consequence of reduced sea-ice (through the albedo feedback), the continued presence of the warm bias in southern summer suggests the possibility that the bias arises in the atmosphere. The decision to control the global temperature with *acllwr* (Section 4.1) preferentially warms cloudy regions and may have contributed. Indeed, simulated downward thermal radiation exhibits a significant bias over the Southern Ocean (Figure 4). A thorough investigation of the source of this bias is beyond the scope of this study, requiring consideration of uncertainties in atmospheric and ocean energy transport, and in solar and thermal radiative transfer, considering clouds, water vapour, and surface processes."

4. Throughout the manuscript, it is stated that the new model is substantially improved from the GENIE-1 (e.g., p.10693, l.15). I think it is very useful to show with figures which part of the simulated climatology is improved.

Many of the improvements are not manifested in the spun-up climatology, but rather reflect the inclusion of dynamics that were previously absent (and forced where necessary), so the improvements are most notably in feedbacks and dynamical variability. One area of significant improvement in climatology is moisture transport and precipitation. Additional text:

"The plotted outputs were chosen to highlight feedbacks that are neglected by the EMBM, viz. 3D dynamical atmospheric transport, providing greatly improved precipitation fields and dynamic surface winds (an imposed forcing in GENIE-1), and interactive clouds (also an imposed forcing field in GENIE-1, comprising a spatiotemporal cloud albedo field and uniform OLR adjustment.)"

and

"The outputs plotted in Figures 3 and 4 were chosen to focus on dynamics that are entirely absent from GENIE-1: interactive winds and interactive clouds. While the inclusion of these dynamics is not expected to improve the simulated climatology (i.e. when compared to simulations that are forced with climatological fields), their inclusion represents a substantial upgrade through the capture of important Earth system feedbacks neglected in GENIE-1. "

and a plot of GENIE-1 vegetative carbon is added for comparison (i.e. 5b, replacing an ENTS soil carbon plot), together with text

"An important example of substantial improvement over the climatology of GENIE-1 is atmospheric moisture transport, previously touched upon in the context of Figure 2. Figure 5 compares PLASIM-GENIE vegetative carbon (5a) and GENIE-1 vegetation carbon (5b, data reproduced from Holden et al, 2013a, Fig 1a) and highlights various aspects of the improved moisture transport. In GENIE-1, deserts are poorly resolved (too moist) and boreal forest does not penetrate far into the continental interior of Eurasia (too dry); these are both shortcomings that arise from diffusive moisture transport (Lenton et al 2006). Although the deserts of the Southern hemisphere are not well resolved in either model, the larger deserts of the Northern hemisphere are distinct in PLASIM-GENIE, while simulated boreal forest penetrates the Eurasian interior."

5. In Figs. 2-4, the model bias, i.e., the difference from the NCEP/NCAR reanalysis should be presented. The model error bar is unknown, otherwise.

Difference plots have been added to Figs 2-4.

Minor comments
1. p.10680, l,22: "that that" should be "that"?
corrected

2. p.10681, l.15: "transport" should be "transfer"? 3. p.10681, l.26: Please explain "self-shading".
Corrected and self-shading explained:

"ENTS includes a parameterisation of self-shading, so that new photosynthetic production is channelled into leaf litter when fractional vegetation coverage approaches 1 and the canopy closes"

4. p.10687, l.15: Would it be more helpful to plot the equation using the revised values of 0.5 and 0.7?
We have addressed this by simplifying the presentation of the equation and introducing some additional text (see response to 1).

5. p.10692, l.15-16: The years of the NCEP/NCAR reanalysis data need to be stated.
6. p.10692, l.16: Which variables are selected? Are these in Table 1? Then, writing "selected variables (Table 1)" is more helpful.
Revised text (comments 5&6):
"Figures 2 to 4 compare a selection of PLASIM-GENIE outputs against NCEP/NCAR reanalysis fields (Kalnay et al 1996). In each case we compare fifty-year PLASIM-GENIE averages of southern summer (JJA) and northern summer (DJF) with the corresponding long-term means (1981-2010) of the reanalysis data. The plotted outputs were chosen to highlight feedbacks neglected by the EMBM, viz. 3D dynamical transport, giving greatly improved precipitation fields and providing dynamic surface winds (an imposed forcing in GENIE-1), and interactive clouds (also an imposed forcing field in GENIE-1, incorporating a spatiotemporal cloud albedo field and uniform OLR adjustment.)"

7. Even if the tuned parameters are mostly ocean model parameters, the description of atmospheric field/circulation and its bias is useful as a description paper of the new coupled model.
A note on atmospheric circulation is added:

"Our focus here is on the wind-stress coupling and the tuned ocean state. The 3D atmospheric circulation is also reasonable. To illustrate, the simulated Southern/Northern hemisphere winter zonal wind jets (~44/43 ms^{-1}, 35°S/35°N, 150mbar) compare to reanalysis data (~41/44 ms^{-1}, 30°S/30°N, 200mbar)."